

LMD



ENTERED

PCT09

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/937,215

TIME: 13:58:25

Input Set : A:\1599-0206P.ST25.txt

Output Set: N:\CRF3\03052002\I937215.raw

3 <110> APPLICANT: Lee, Se Yong et al.
 5 <120> TITLE OF INVENTION: Trehalose Synthase Protein, Gene, Plasmids, Microorganisms,
 and A Process

6 for Producing Trehalose
 8 <130> FILE REFERENCE: 1599-0206P
 10 <140> CURRENT APPLICATION NUMBER: US 09/937,215
 11 <141> CURRENT FILING DATE: 2001-09-24
 13 <160> NUMBER OF SEQ ID NOS: 6
 15 <170> SOFTWARE: PatentIn version 3.1
 17 <210> SEQ ID NO: 1
 18 <211> LENGTH: 4753
 19 <212> TYPE: DNA
 20 <213> ORGANISM: Pseudomonas stutzeri
 22 <220> FEATURE:
 23 <221> NAME/KEY: CDS
 24 <222> LOCATION: (1537)..(3603)
 25 <223> OTHER INFORMATION:
 28 <400> SEQUENCE: 1

```

29 gatcgctggc gtactgcagg tagagcaggc gcatcggcc ccagggcgca tcggccggct      60
31 ccgctgtgcc ctgctgggtc atgaagcgga cgaagcggcc atcgcggaac cgtggacgcc      120
33 attcggggct gtcggggctc cggctgtcgg tgagcgtgcg ccacaggtcg ctgcgaaacg      180
35 gcggaccgct ccaaagcgcg ccgtggatgg gatcgccgag cagttcgtgc agctcccagg      240
37 aacgttgcca atgcagcgcg ccgaggctca ggccatgcag atacaggcgc ggtcggcggt      300
39 cggccggcag ttcggtccag tagccataga tctcggcgaa tagcgcgcgg gccacgtcgc      360
41 ggccgtagtc ggcctccacc agcagcgcca gcgggctggt cagataggag tactgcaacg      420
43 ccacgctggc gatatgcgcg tgggtgcagg attccactgc gttcatcgcc gccgggtcga      480
45 tccagccggt accggtgggc gtcaccagca ccagcaccga tcgctcgaag gcgccgctgc      540
47 gctgcagctc gcgcaaggcc agacgcgcgc gctggcgcgg ggtctctgcc gcgcgcagac      600
49 cgacgtagac gcgaatcggc tcgagcgccg agcggccgct caagacgctg atatccgccc      660
51 ccgacgggcc ggagccgatg aactcgcggc cgggtgcggc cagctcctcc cagcgcagca      720
53 acgaggcccg gctgccgctt ttcagcgcgg aggcgggtgg cgccgtctcc ggttcgatca      780
55 gggcgctgta ctgcgcgaag gatgcgtcca gcatgcgcag tgcccgcgcc gccagcacat      840
57 cgctgagcag cgaccagaac agcgcacgcy ccaccagcac gccgatcacg ttggccaggc      900
59 gccgtggcag cagcgggtcg gcgtgccgcg agacgaagcg cgacaccagc cgatacagac      960
61 gcgccagcgt cagcaggatg agaaaggctc ccagcgcggt gagaatgact tcgagcagggt      1020
63 gcgcaactgt caccggcggc atgcccatac gcgcgcgtac cgcgttctgc cagccggcga      1080
65 cctggtctag gaaatacccg gccagcagca ggcagccgac cgcgatcagc agattgaccc      1140
67 gctcgcgctg ccagcctggg cgtcccgcca gttccagata gcgccacagc cagcgcagca      1200
69 acacgccgag gccatagccc accgccagcg ccgcgcgggc cagcacgccc tggctcagcg      1260
71 tcgagcgcgg cagcagcgat ggcgtcagcg ccgcgcagaa gaacagcgtg cccagcagca      1320
73 ggccgaaacc ggacagcagc cgccagatat agaggacggg caggtgcagc atgaagatct      1380
75 ccgcggctcg gtgacggcgt cgcgcctcgg catctcaggg cgtgtccggg cgtgcgggtc      1440
77 ccgtgatggt ccgcagcagg ccaatccgat gcaacgatgg ccgagcggcc gactcaaacg      1500
79 totacatttc cctagtgtcg ccggaaccga tcgccg atg agc atc cca gac aac      1554

```

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/937,215

TIME: 13:58:25

Input Set : A:\1599-0206P.ST25.txt

Output Set: N:\CRF3\03052002\I937215.raw

80		Met Ser Ile Pro Asp Asn	
81		1 5	
83	acc tat atc gaa tgg ctg gtc agc cag tcc atg ctg cat gcg gcc cgc	1602	
84	Thr Tyr Ile Glu Trp Leu Val Ser Gln Ser Met Leu His Ala Ala Arg		
85	10 15 20		
87	gag cgg tcg cgt cat tac gcc ggc cag gcg cgt ctc tgg cag cgg cct	1650	
88	Glu Arg Ser Arg His Tyr Ala Gly Gln Ala Arg Leu Trp Gln Arg Pro		
89	25 30 35		
91	tat gcc cag gcc cgc ccg cgc gat gcc agc gcc atc gcc tcg gtg tgg	1698	
92	Tyr Ala Gln Ala Arg Pro Arg Asp Ala Ser Ala Ile Ala Ser Val Trp		
93	40 45 50		
95	ttc acc gcc tat ccg gcg gcc atc atc acg ccg gaa ggc ggc acg gta	1746	
96	Phe Thr Ala Tyr Pro Ala Ala Ile Ile Thr Pro Glu Gly Gly Thr Val		
97	55 60 65 70		
99	ctc gag gcc ctc ggc gac gac cgc ctc tgg agt gcg ctc tcc gaa ctc	1794	
100	Leu Glu Ala Leu Gly Asp Asp Arg Leu Trp Ser Ala Leu Ser Glu Leu		
101	75 80 85		
103	ggc gtg cag ggc atc cac aac ggg ccg atg aag cgt tcc ggt ggc ctg	1842	
104	Gly Val Gln Gly Ile His Asn Gly Pro Met Lys Arg Ser Gly Gly Leu		
105	90 95 100		
107	cgc gga cgc gag ttc acc ccg acc atc gac ggc aac ttc gac cgc atc	1890	
108	Arg Gly Arg Glu Phe Thr Pro Thr Ile Asp Gly Asn Phe Asp Arg Ile		
109	105 110 115		
111	agc ttc gat atc gac ccg agc ctg ggg acc gag gag cag atg ctg cag	1938	
112	Ser Phe Asp Ile Asp Pro Ser Leu Gly Thr Glu Glu Gln Met Leu Gln		
113	120 125 130		
115	ctc agc cgg gtg gcc gcg gcg cac aac gcc atc gtc atc gac gac atc	1986	
116	Leu Ser Arg Val Ala Ala His Asn Ala Ile Val Ile Asp Asp Ile		
117	135 140 145 150		
119	gtg ccg gca cac acc ggc aag ggt gcc gac ttc cgc ctc gcg gaa atg	2034	
120	Val Pro Ala His Thr Gly Lys Gly Ala Asp Phe Arg Leu Ala Glu Met		
121	155 160 165		
123	gcc tat ggc gac tac ccc ggg ctg tac cac atg gtg gaa atc cgc gag	2082	
124	Ala Tyr Gly Asp Tyr Pro Gly Leu Tyr His Met Val Glu Ile Arg Glu		
125	170 175 180		
127	gag gac tgg gag ctg ctg ccc gag gtg ccg gcc ggg cgt gat tcg gtc	2130	
128	Glu Asp Trp Glu Leu Leu Pro Glu Val Pro Ala Gly Arg Asp Ser Val		
129	185 190 195		
131	aac ctg ctg ccg ccg gtg gtc gac ccg ctc aag gaa aag cac tac atc	2178	
132	Asn Leu Leu Pro Pro Val Val Asp Arg Leu Lys Glu Lys His Tyr Ile		
133	200 205 210		
135	gtc gcc cag ctg cag ccg gtg atc ttc ttc gag ccg gcc atc aag gac	2226	
136	Val Gly Gln Leu Gln Arg Val Ile Phe Phe Glu Pro Gly Ile Lys Asp		
137	215 220 225 230		
139	acc gac tgg agc gtc acc ggc gag gtc acc ggg gtc gac ggc aag gtg	2274	
140	Thr Asp Trp Ser Val Thr Gly Glu Val Thr Gly Val Asp Gly Lys Val		
141	235 240 245		
143	cgt cgc tgg gtc tat ctg cac tac ttc aag gag ggc cag ccg tcg ctg	2322	
144	Arg Arg Trp Val Tyr Leu His Tyr Phe Lys Glu Gly Gln Pro Ser Leu		

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/937,215

TIME: 13:58:25

Input Set : A:\1599-0206P.ST25.txt

Output Set: N:\CRF3\03052002\I937215.raw

145	250	255	260	
147 aac tgg ctc gac ccg acc ttc gcc gcg cag cag ctg atc atc ggc gat				2370
148 Asn Trp Leu Asp Pro Thr Phe Ala Ala Gln Gln Leu Ile Ile Gly Asp				
149	265	270	275	
151 gcg ctg cac gcc atc gac gtc acc ggc gcc cgg gtg ctg cgc ctg gac				2418
152 Ala Leu His Ala Ile Asp Val Thr Gly Ala Arg Val Leu Arg Leu Asp				
153	280	285	290	
155 gcc aac ggc ttc ctc ggc gtg gaa cgg cgc gcc gag ggc acg gcc tgg				2466
156 Ala Asn Gly Phe Leu Gly Val Glu Arg Arg Ala Glu Gly Thr Ala Trp				
157 295	300	305	310	
159 tcg gag ggc cac ccg ctg tcc gtc acc ggc aac cag ctg ctc gcc ggg				2514
160 Ser Glu Gly His Pro Leu Ser Val Thr Gly Asn Gln Leu Leu Ala Gly				
161	315	320	325	
163 gcg atc cgc aag gcc ggc ggc ttc agc ttc cag gag ctg aac ctg acc				2562
164 Ala Ile Arg Lys Ala Gly Gly Phe Ser Phe Gln Glu Leu Asn Leu Thr				
165	330	335	340	
167 atc gat gac atc gcc gcc atg tcc cac ggc ggg gcc gat ctg tcc tac				2610
168 Ile Asp Asp Ile Ala Ala Met Ser His Gly Gly Ala Asp Leu Ser Tyr				
169	345	350	355	
171 gac ttc atc acc cgc ccg gcc tat cac cat gcg ttg ctc acc ggc gat				2658
172 Asp Phe Ile Thr Arg Pro Ala Tyr His His Ala Leu Leu Thr Gly Asp				
173	360	365	370	
175 acc gaa ttc ctg cgc atg atg ctg cgc gaa gtg cac gcc ttc ggc atc				2706
176 Thr Glu Phe Leu Arg Met Met Leu Arg Glu Val His Ala Phe Gly Ile				
177 375	380	385	390	
179 gac ccg gcg tca ctg atc cat gcg ctg cag aac cat gac gag ttc acc				2754
180 Asp Pro Ala Ser Leu Ile His Ala Leu Gln Asn His Asp Glu Phe Thr				
181	395	400	405	
183 ctg gag ctg gtg cac ttc tgg acg ctg cac gcc tac gac cat tac cac				2802
184 Leu Glu Leu Val His Phe Trp Thr Leu His Ala Tyr Asp His Tyr His				
185	410	415	420	
187 tac aag ggc cag acc ctg ccc ggc ggc cac ctg cgc gaa cat atc cgc				2850
188 Tyr Lys Gly Gln Thr Leu Pro Gly Gly His Leu Arg Glu His Ile Arg				
189	425	430	435	
191 gag gaa atg tac gag cgg ctg acc ggc gaa cac gcg ccg tac aac ctc				2898
192 Glu Glu Met Tyr Glu Arg Leu Thr Gly Glu His Ala Pro Tyr Asn Leu				
193	440	445	450	
195 aag ttc gtc acc aac ggg gtg tcc tgc acc acc gcc agc gtg atc gcc				2946
196 Lys Phe Val Thr Asn Gly Val Ser Cys Thr Thr Ala Ser Val Ile Ala				
197 455	460	465	470	
199 gcg gcg ctt aac atc cgt gat ctg gac gcc atc ggc ccg gcc gag gtg				2994
200 Ala Ala Leu Asn Ile Arg Asp Leu Asp Ala Ile Gly Pro Ala Glu Val				
201	475	480	485	
203 gag cag atc cag cgt ctg cat atc ctg ctg gtg atg ttc aat gcc atg				3042
204 Glu Gln Ile Gln Arg Leu His Ile Leu Leu Val Met Phe Asn Ala Met				
205	490	495	500	
207 cag ccc ggc gtg ttc gcc ctc tcc ggc tgg gat ctg gtc ggc gcc ctg				3090
208 Gln Pro Gly Val Phe Ala Leu Ser Gly Trp Asp Leu Val Gly Ala Leu				
209	505	510	515	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/937,215

DATE: 03/05/2002

TIME: 13:58:25

Input Set : A:\1599-0206P.ST25.txt

Output Set: N:\CRF3\03052002\I937215.raw

211	ccg	ctg	gcg	ccc	gag	cag	gtc	gag	cac	ctg	atg	ggc	gat	ggc	gat	acc	3138
212	Pro	Leu	Ala	Pro	Glu	Gln	Val	Glu	His	Leu	Met	Gly	Asp	Gly	Asp	Thr	
213		520					525					530					
215	cgc	tgg	atc	aat	cgc	ggc	ggc	tat	gac	ctc	gcc	gat	ctg	gcg	ccg	gag	3186
216	Arg	Trp	Ile	Asn	Arg	Gly	Gly	Tyr	Asp	Leu	Ala	Asp	Leu	Ala	Pro	Glu	
217	535					540					545				550		
219	gcg	tcg	gtc	tcc	gcc	gaa	ggc	ctg	ccc	aag	gcc	cgc	tcg	ctg	tac	ggc	3234
220	Ala	Ser	Val	Ser	Ala	Glu	Gly	Leu	Pro	Lys	Ala	Arg	Ser	Leu	Tyr	Gly	
221				555						560					565		
223	agc	ctg	gcc	gag	cag	ctg	cag	cgg	cca	ggc	tcc	ttc	gcc	tgc	cag	ctc	3282
224	Ser	Leu	Ala	Glu	Gln	Leu	Gln	Arg	Pro	Gly	Ser	Phe	Ala	Cys	Gln	Leu	
225				570					575					580			
227	aag	cgc	atc	ctc	agc	gtg	cgc	cag	gcc	tac	gac	atc	gct	gcc	agc	aag	3330
228	Lys	Arg	Ile	Leu	Ser	Val	Arg	Gln	Ala	Tyr	Asp	Ile	Ala	Ala	Ser	Lys	
229		585					590					595					
231	cag	atc	ctg	att	ccg	gat	gtg	cag	gcg	ccg	gga	ctc	ctg	gtg	atg	gtc	3378
232	Gln	Ile	Leu	Ile	Pro	Asp	Val	Gln	Ala	Pro	Gly	Leu	Leu	Val	Met	Val	
233		600					605					610					
235	cac	gag	ctg	cct	gcc	ggc	aag	ggc	gtg	cag	ctc	acg	gca	ctg	aac	ttc	3426
236	His	Glu	Leu	Pro	Ala	Gly	Lys	Gly	Val	Gln	Leu	Thr	Ala	Leu	Asn	Phe	
237	615					620					625				630		
239	agc	gcc	gag	ccg	gtc	agc	gag	acc	atc	tgc	ctg	ccc	ggc	gtg	gcg	ccc	3474
240	Ser	Ala	Glu	Pro	Val	Ser	Glu	Thr	Ile	Cys	Leu	Pro	Gly	Val	Ala	Pro	
241				635					640					645			
243	ggc	ccg	gtg	gtg	gac	atc	att	cac	gag	agt	gtg	gag	ggc	gac	ctc	acc	3522
244	Gly	Pro	Val	Val	Asp	Ile	Ile	His	Glu	Ser	Val	Glu	Gly	Asp	Leu	Thr	
245				650					655					660			
247	gac	aac	tgc	gag	ctg	cag	atc	aac	ctc	gac	ccg	tac	gag	ggg	ctt	gcc	3570
248	Asp	Asn	Cys	Glu	Leu	Gln	Ile	Asn	Leu	Asp	Pro	Tyr	Glu	Gly	Leu	Ala	
249		665					670					675					
251	ctg	cgt	gtg	gtg	agc	gcc	gcg	ccg	ccg	gtg	atc	tgagcgcct	cttcgcgcgc				3623
252	Leu	Arg	Val	Val	Ser	Ala	Ala	Pro	Pro	Val	Ile						
253		680					685										
255	cccggttcg	ccgctatagt	gcgcagcgcc	tggggcgcgc	attgccctcg	ccgtcgagac											3683
257	cagcccggtgt	cggtcacttc	gcttttcgcg	cttgcgctgc	tgcgcgtggc	gctgcttgcc											3743
259	gcacccgtct	gggcgcagac	cgccgtcccg	cccgccagc	agccgatctg	cctgagcggc											3803
261	agctgctct	gcgtgccggc	cgccgccagc	gatccacagg	cggtctacga	ccgcgtccag											3863
263	cgtatggcta	cgctggccct	gcagaactgg	atccagcagt	cgcgcgaccg	cctgatggcc											3923
265	ggcggcgctg	agccgatacc	gctgcacatc	cgctcgcagc	tcgagccgta	tttcgatctt											3983
267	gccgtgctgg	agagtgcgcy	gtaccgcgtc	ggcgacgagg	tggtgctgac	tgccggcaac											4043
269	accctgctgc	gcaaccggga	cgtcaatgcc	gtgaccctga	tcgacgtcat	cgtcttccgc											4103
271	cacgaggagg	atgcccggga	caacgtcgcg	ctctggggcc	atgagctcaa	gcacgtcgag											4163
273	caatatctgg	actggggcgt	cgccgagttc	gcccggcgct	atacgcagga	tttccgtgcc											4223
275	gtggagcgcc	cgccctatgc	gctggagcgt	gaggtggaag	aggccctgcg	cgagacgcag											4283
277	acgcggcgct	gagcgagctg	atcggtgctg	ctgcccgcac	tggtgctgaag	cccaccaatg											4343
279	acgcggcgca	aaacgaaaaa	ccccgccgag	gcggggtttc	tgacgcgggt	tgtgcggtca											4403
281	gctcagaacg	ccgggaccac	ggcgcccttg	tacttttctc	cgatgaactg	gcgtacttgc											4463
283	tcgctgtgca	gcgcggcagc	cagtttctgc	atggcatcgc	tgctcttggt	gtccggacgg											4523
285	gcgaccagaa	tgttcacgta	tggcgagtcg	ctgccctcga	tcaccagggc	gtcctgggtc											4583

RAW SEQUENCE LISTING

DATE: 03/05/2002

PATENT APPLICATION: US/09/937,215

TIME: 13:58:25

Input Set : A:\1599-0206P.ST25.txt

Output Set: N:\CRF3\03052002\I937215.raw

```

287 ggggttcagct tggcttccag cgcgtagttg gtgttgatca gcgccaggtc gacctggggtc 4643
289 agcacgcgcg gcagagtcgc ggcttccagt tcgcggatct tgatcttctt cgggttctcg 4703
291 gcgatgtctt cggcgtggcg gtgatgccgg cgcgctcctt cagaccgatc 4753
294 <210> SEQ ID NO: 2
295 <211> LENGTH: 689
296 <212> TYPE: PRT
297 <213> ORGANISM: Pseudomonas stutzeri
299 <400> SEQUENCE: 2
301 Met Ser Ile Pro Asp Asn Thr Tyr Ile Glu Trp Leu Val Ser Gln Ser
302 1 5 10 15
305 Met Leu His Ala Ala Arg Glu Arg Ser Arg His Tyr Ala Gly Gln Ala
306 20 25 30
309 Arg Leu Trp Gln Arg Pro Tyr Ala Gln Ala Arg Pro Arg Asp Ala Ser
310 35 40 45
313 Ala Ile Ala Ser Val Trp Phe Thr Ala Tyr Pro Ala Ala Ile Ile Thr
314 50 55 60
317 Pro Glu Gly Gly Thr Val Leu Glu Ala Leu Gly Asp Asp Arg Leu Trp
318 65 70 75 80
321 Ser Ala Leu Ser Glu Leu Gly Val Gln Gly Ile His Asn Gly Pro Met
322 85 90 95
325 Lys Arg Ser Gly Gly Leu Arg Gly Arg Glu Phe Thr Pro Thr Ile Asp
326 100 105 110
329 Gly Asn Phe Asp Arg Ile Ser Phe Asp Ile Asp Pro Ser Leu Gly Thr
330 115 120 125
333 Glu Glu Gln Met Leu Gln Leu Ser Arg Val Ala Ala His Asn Ala
334 130 135 140
337 Ile Val Ile Asp Asp Ile Val Pro Ala His Thr Gly Lys Gly Ala Asp
338 145 150 155 160
341 Phe Arg Leu Ala Glu Met Ala Tyr Gly Asp Tyr Pro Gly Leu Tyr His
342 165 170 175
345 Met Val Glu Ile Arg Glu Glu Asp Trp Glu Leu Leu Pro Glu Val Pro
346 180 185 190
349 Ala Gly Arg Asp Ser Val Asn Leu Leu Pro Pro Val Val Asp Arg Leu
350 195 200 205
353 Lys Glu Lys His Tyr Ile Val Gly Gln Leu Gln Arg Val Ile Phe Phe
354 210 215 220
357 Glu Pro Gly Ile Lys Asp Thr Asp Trp Ser Val Thr Gly Glu Val Thr
358 225 230 235 240
361 Gly Val Asp Gly Lys Val Arg Arg Trp Val Tyr Leu His Tyr Phe Lys
362 245 250 255
365 Glu Gly Gln Pro Ser Leu Asn Trp Leu Asp Pro Thr Phe Ala Ala Gln
366 260 265 270
369 Gln Leu Ile Ile Gly Asp Ala Leu His Ala Ile Asp Val Thr Gly Ala
370 275 280 285
373 Arg Val Leu Arg Leu Asp Ala Asn Gly Phe Leu Gly Val Glu Arg Arg
374 290 295 300
377 Ala Glu Gly Thr Ala Trp Ser Glu Gly His Pro Leu Ser Val Thr Gly
378 305 310 315 320
381 Asn Gln Leu Leu Ala Gly Ala Ile Arg Lys Ala Gly Gly Phe Ser Phe

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/937,215

DATE: 03/05/2002

TIME: 13:58:26

Input Set : A:\1599-0206P.ST25.txt

Output Set: N:\CRF3\03052002\I937215.raw

09937215-099401